ROOT CAUSE ANALYSIS OF CYBERSECURITY INCIDENTS ON PIPELINES USING THE NFR APPROACH

NARY SUBRAMANIAN

Department of Computer Science
The University of Texas at Tyler
3900 University Blvd, Tyler, TX 75799, USA

(received: 23 June 2021; revised: 15 July 2021;
accepted: 10 August 2021; published online: 30 November 2021)

Abstract: Pipelines transporting oil, gas, water, and other substances form part of the critical infrastructure of the society and are mostly controlled by advanced automation technology. This automation enables remote control and monitoring of pipeline operations by means of wide area networks that include microwaves, satellites, and cellular technologies. Often these pipeline control systems are also connected to the Internet to permit their operational control from anywhere. However, this bridging of the so-called "air-gap" between the critical infrastructure control system and the Internet has also introduced cybersecurity weaknesses that allows malicious actors to take control away from legitimate users of the system. While cybersecurity needs to be built into the system during the design phase itself, it is important, especially after a cybersecurity incident, to know the actual causes behind the incident so that appropriate countermeasures may be taken quickly to avoid a recurrence of the incident. Typical techniques to identify these root causes include five whys, fishbone diagrams, and causal factors analysis; this paper presents an alternate technique to identify root causes for pipeline cybersecurity incidents based on the NFR Approach where NFR stands for Non-Functional Requirements of the pipeline system. The NFR Approach starts with the requirements for the system in the first place, establishes the relationships between the design of the system and its requirements, and then identifies the root causes in a structured manner. In this paper, the NFR Approach is applied to analyze root causes of the Florida water system attack that occurred in February 2021. The advantages of the NFR Approach over traditional methods to identify root causes especially for pipeline incidents include the traceability of the causes to the requirements of the system, identification of synergistic and conflicting operational goals, and historical record-keeping.

Keywords: root cause, cybersecurity, pipeline, critical infrastructure, NFR Approach
DOI: https://doi.org/10.17466/tq2021/25.3/b

1. Introduction

Pipelines [1] are used to transport critical fluids such as natural gas, gasoline, chemicals, water, and others over long distances so that they can reach

consuming centers from the production or distribution centers. For example, oil and gas pipelines traverse the continental United States for thousands of miles. As such, these pipelines become part of the critical infrastructure of the nation and require special attention to their cybersecurity [2–4] needs since most of these pipelines are controlled by advanced automation [5]. This automation enables remote control and monitoring of pipeline operation by means of wide area networks that include microwaves, satellites, and cellular technologies. However, several cybersecurity attacks on pipeline systems including the water supply system in Florida [6] in February, 2021, and the Colonial pipeline [7] in May, 2021, show the importance of a secure system for controlling and monitoring pipeline operations; in the attack on the water supply system in Florida, the attacker was able to change the chemicals in the water while in the Colonial pipeline attack, the system was disabled by ransomware. There have been several other pipeline incidents as well [8].

Cybersecurity is the provision of measures to ensure that a system is secure from digital attacks so that confidentiality, integrity, and availability of the system are preserved [9]. Since a system often includes hardware, software, networks, people, policies, and procedures, cybersecurity must ensure that the system is not vulnerable from any of these components. For example, hardware access will allow external entities to extract data from the system memory directly, software access will allow external entities to monitor activity in the system, network access will allow external entities to monitor and extract network data, access to people will permit social engineering attacks to get information about the system, knowledge of policies will allow understanding of the parameters typically used for system access and use including the lengths of passwords, and understanding of procedures will allow external entities to extract information by following operating procedures of the company such as, for example, resetting the password. Measures to improve cybersecurity include addition of new technology, human resource training, and enhanced policies and procedures.

Often pipeline control systems are connected to the Internet to permit their operational control from anywhere. This bridging of the so-called “air-gap” between the critical infrastructure control system and the Internet has also introduced cybersecurity weaknesses that allows malicious actors to take control away from the legitimate users of the system. While cybersecurity needs to be built into the system during the design phase itself, it is important, especially after a cybersecurity incident, to know the actual causes behind the incident so that appropriate countermeasures may be taken quickly to avoid a recurrence of the incident. This is because the majority of cybersecurity incidents occur due to non-technical factors and therefore purely technology-based solutions may not be sufficient for providing cybersecurity protection. Moreover, most of these non-technical factors are usually related to human error [10], so understanding true causes, also called root causes, for cybersecurity breaches will help prevent a recurrence.

Typical techniques to identify root causes include [11, 12] five whys, fishbone diagrams, and causal factors analysis; in the five whys technique, to determine the root cause, for each problem the probable cause for the problem (the first "why") is questioned and for each answer to this question, repeatedly "why" is asked for up to four more times. In the fishbone diagram technique, also called Ishikawa diagram or cause-and-effect diagram, the problem is stated at the head of the fish and the backbone is a line joining the head; categories of possible causes are written along the ribs that connect to the backbone and for each category, potential causes are written as smaller bones perpendicular to the ribs. In the causal factors analysis technique, root causes are identified by conducting a brainstorming session with all stakeholders where answers to questions such as what could have caused the incident are explored. These techniques are generic enough to apply to any domain including pipeline cybersecurity incidents. Also a study of literature [13, 14] reveals that usual analyses of cybersecurity incidents are from a data-analytic viewpoint wherein researchers have tried to identify causes for incidents using statistical techniques.

This paper presents an alternate technique to identify root causes for pipeline cybersecurity incidents based on the NFR Approach [15] where NFR stands for Non-Functional Requirements of the pipeline system including security, safety, reliability, maintainability, and others. The NFR Approach starts with the requirements for the system in the first place and then identifies the root causes in a structured manner. In this paper, the NFR Approach's ability to analyze cybersecurity incidents on pipelines is demonstrated by applying it to study the Florida water system attack [6] that occurred in February, 2021. The advantages of the NFR Approach over traditional methods to identify root causes especially for pipeline incidents, include the traceability of the causes to the requirements of the system, identification of synergistic and conflicting operational goals, and historical record-keeping.

This paper is organized as follows: Section 2 discusses the architecture of a typical pipeline system, Section 3 introduces the NFR Approach for identifying root causes, Section 4 discusses the application of the NFR Approach to the Florida water system incident, Section 5 discusses the results, and Section 6 concludes the paper.

2. Architecture of a Pipeline System

Figure 1 shows the architecture of a typical pipeline system [4, 5]. Long stretches of pipelines connect source to destination; source may be gas production facility, water supply station, or chemical manufacturer, and destination is the place where these fluids are consumed. However, these fluids lose speed as they travel from source to destination and to ensure they are traveling at the required speed, periodically, along the length of the pipeline there are pumping stations that increase the pressure of the fluids on the pipes; these pumping stations also help divert fluids as required along different pipes at junctions. Also along

the length of the pipeline there are monitoring stations that measure the speed of fluids, their pressure, and other physical properties appropriate to the fluids and transmit these data to the control center; these monitoring stations may be coincident with pumping stations or separate from them.

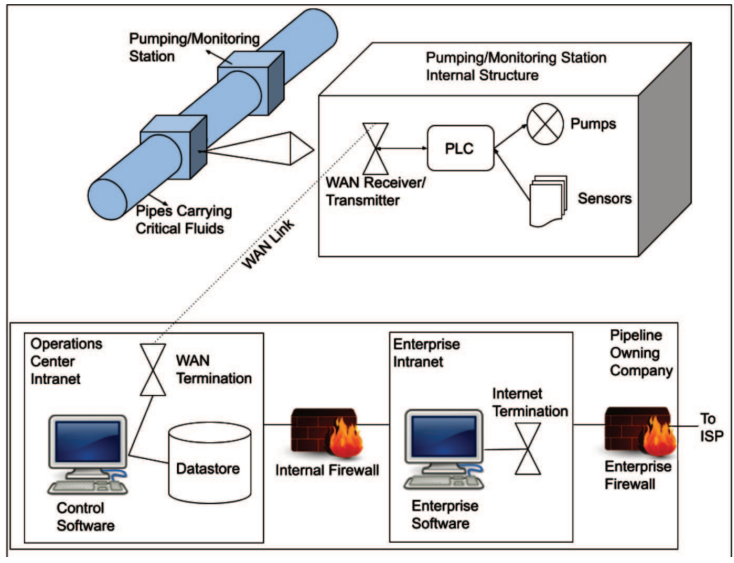

Figure 1. Architecture of a typical pipeline system

Figure 1 also shows the structure of a typical pumping station: the Programmable Logic Controller (PLC) is the main component of the station that has the necessary logic to drive the pumps as well as read any measurement sensors in the station. The stimulus for pumping may be received by the PLC from the supervisory control at the operations center of the company owning the pipeline; as soon as the pumps start, the pressure of the fluid in the pipes connected to the pumps increases and the fluid speed increases. In a similar manner the stimulus for measuring the physical parameters of the fluid flowing through the pipes may be received from the supervisory control station and the data from the sensors are sent to the control station. The connection between the PLC and the operations center is typically over a wide area network (WAN) link provided usually by third-party services including satellite links, cellular telephony, microwaves, and high-speed wired links.

The operations center at the pipeline owning company has a bank of computers at which operators control and monitor the entire pipeline network. The control software that runs on these computers provides the operators with a real-time view of the entire network. In case of any change in the physical parameters of the fluids being transported over the pipes, appropriate actions may be remotely taken from the control station by sending commands to the PLC's at the pumping and monitoring stations distributed throughout the network. All

commands sent to the pumping and monitoring stations are logged in the data store for historical reference and any accident audits by regulators. Also, all data received from monitoring stations are collected by the data acquisition system and then stored in the data store. The operations center is also connected to the WAN to which all pumping and monitoring stations are connected. Frequently, all elements in the operations center are collectively referred to as the control system. The intranet in the operations center is connected to the enterprise intranet through an internal firewall - this is the so-called "air-gap" in the critical infrastructure control. The enterprise intranet hosts enterprise software, database, hardware, and users that use these systems. The enterprise intranet is also connected to the ISP (Internet Services Provider) through the enterprise firewall for Internet access.

3. The NFR Approach

The NFR Approach [15] is a goal-oriented approach that can be applied to determine the extent to which objectives are achieved by a design – here the objectives are defined as identifying the root causes for a cybersecurity incident in a pipeline control system. NFR stands for Non-Functional Requirements that are properties of a system such as security, reliability, maintainability, flexibility, human factors, supportability, or scalability, and could equally well represent functional objectives and constraints for a system. The NFR Approach uses a well-defined ontology for this purpose that includes NFR softgoals, operationalizing softgoals, cause softgoals, claim softgoals, contributions, tracebacks, and propagation rules; each of these elements is described briefly below. Furthermore, since strictly quantitative assessment of soft or vaguely defined properties is difficult, the NFR Approach uses the concept of *satisficing*, a term borrowed from economics, which indicates satisfaction within limits instead of absolute satisfaction of the goal.

NFR softgoals represent NFR's and their decompositions. Elements that have physical equivalents (process, product, or design elements) are represented by operationalizing softgoals and their decompositions. Each cause for the cybersecurity incident is captured by a cause softgoal. Each softgoal is named using the convention

$$Type[Topic1,\ Topic2,\ \ldots]$$

where $Type$ is the name of the softgoal and $Topic$ (could be zero or more) is the context where the softgoal is used. Topic is optional for a softgoal; for a claim softgoal, which is a softgoal capturing a design or organizational decision, the name may be the rationale itself. Softgoals may be decomposed into other softgoals in three ways: in an AND-contribution, satisficing all child softgoals is essential to satisfice the parent; in an OR-contribution, satisficing one child softgoal is sufficient to satisfice the parent; in a refinement, a parent has only one child and the parent is satisficed if the child is satisficed.

Contributions (MAKE, HELP, HURT, and BREAK) are usually made between softgoals: between NFR softgoals and other NFR softgoals, between operationalizing softgoals and NFR softgoals, between operationalizing softgoals and other operationalizing softgoals, between cause softgoals and operationalizing softgoals, between cause softgoals themselves, and between claim softgoals and other elements in the SIG. Reasons for these contributions are captured by claim softgoals and, in this case, there is a contribution between a claim softgoal and the contribution being justified. Each of the four types of contributions has a specific semantic significance: MAKE contribution refers to a strongly positive degree of satisficing of objectives by artifacts (could be design decisions as well) under consideration, HELP contribution refers to a positive degree of satisficing, HURT contribution refers to a negative degree of satisficing, and BREAK contribution refers to a strongly negative degree of satisficing. The partial ontology of the NFR Approach is shown in Figure 2.

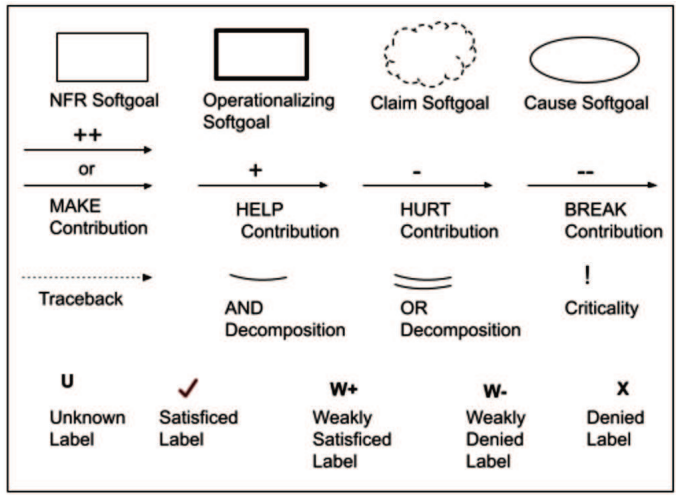

Figure 2. The ontology of the NFR approach

Tracebacks are contributions between cause softgoals and operationalizing softgoals that indicate the design element(s) responsible for those cause softgoals and capture the potential root causes for the cybersecurity incident. Due to MAKE, HELP, HURT, and BREAK contributions, some of the softgoals, decompositions, tracebacks, and contributions themselves acquire labels that capture the extent to which that element is satisficed: satisficed, weakly satisficed, weakly denied (or weakly not satisficed), denied (or not satisficed), or unknown (in the ontology shown in Figure 2, unknown labels are marked by 'U'; however, to avoid clutter, often this label is omitted for elements whose labels are not yet known). Moreover, high priority softgoals, decompositions, tracebacks, and contributions may be indicated using the criticality symbol. The graph that captures the softgoals, their decompositions, their contributions, and the tracebacks is called the

Softgoal Interdependency Graph (SIG). Propagation rules propagate labels from a child softgoal to its parent across decompositions and across contributions; propagation rules aid in the rationalization process of the NFR Approach. Propagation rules of relevance to the discussion in this paper include (details can be seen in [15]):

R1. Determine labels for all NFR softgoals, operationalizing softgoals, cause softgoals, claim softgoals, contributions, and tracebacks: each is either satisficed, denied, weakly satisficed, weakly denied, or unknown.

R2. If a softgoal label is satisficed (denied) and it has a MAKE-contribution to its parent, then the softgoal propagates its label to the parent.

R3. If a softgoal label is satisficed (denied) and it has a BREAK-contribution to its parent, then the softgoal propagates denied (satisficing) label to its parent.

R4. If all labels propagated to a parent (either softgoal, contribution, or traceback) are satisficed, then that parent is satisficed.

R5. If all labels propagated to a parent (either softgoal, contribution, or traceback) are denied, then that parent is denied.

R6. If a traceback is satisficed, it propagates the label of the child to the parent.

R7. In the case of AND-decomposed softgoals, if even one child softgoal has a denied label then the parent is denied; else the parent is satisficed.

R8. In the case of OR-decomposed softgoals, if even one child softgoal has a satisficed label then the parent is satisficed; else the parent is denied.

R9. If the parent softgoal has the label X (where X is one of satisficed, weakly satisficed, weakly denied, denied, or unknown) and the parent is decomposed into children with an AND-contribution, then at least one child has the same label X where X is the least significant label in the order satisficed > weakly satisficed > unknown > weakly denied > denied.

R10. If the parent softgoal has the label X (where X is one of satisficed, weakly satisficed, weakly denied, denied, or unknown) and the parent is decomposed into children with an OR-contribution, then at least one child has the same label X where X is the most significant label in the order satisficed > weakly satisficed > unknown > weakly denied > denied.

The propagation rule R1 states that a softgoal can have one of five labels – satisficed, weakly satisficed, weakly denied, denied, and unknown. Rules R2 and R3 state the labels propagated by a softgoal to its parent via MAKE or BREAK contributions - across a MAKE it remains the same while across a BREAK it inverts. Rules R4 and R5 state the labels for parents based on contributions from their children – if all labels are satisficed, then the parent is satisficed and if all labels are denied then the parent is denied. Rule R6 states how a traceback propagates the child's label to the parent when the traceback is satisficed. Rules R7 and R8 state the labels propagated to the parent softgoal involved in an

AND- or OR-decomposition with its children. Rules R9 and R10 describe the inverse relationship to R7 and R8: R9 and R10 state the expectation from children when the labels for the parents are known; R9 says that if the parent in an AND-decomposition is of a specific label then that label should be the least label of its children where the ordering goes as satisficed > weakly satisficed > unknown > weakly denied > denied. Likewise, R10 says that if the parent in an OR-decomposition is of a specific label then that label should be the same as the child with the highest label where the ordering goes as satisficed > weakly satisficed > unknown > weakly denied > denied. There are eight iterative steps for applying the NFR Approach for identifying root causes for a cybersecurity incident in pipelines:

1. Decompose the cybersecurity requirements for the pipeline system into NFR softgoals.
2. Decompose the architecture of the pipeline system into its constituent operationalizing softgoals.
3. Determine the contributions made by the operationalizing softgoals to the NFR softgoals.
4. Analyze the cybersecurity incident report to identify causes and associate them with corresponding cause softgoals; then capture the contributions between these cause softgoals and operationalizing softgoals.
5. Using the cybersecurity body of knowledge, identify causes for existing cause softgoals - that is derive a cause softgoal chain.
6. Associate tracebacks with the leaf cause softgoals to the corresponding operationalizing softgoals.
7. Identify the root causes from the tracebacks by applying the propagation rules of the NFR Approach.
8. Capture justifications for all elements in the SIG by means of claim softgoals.

In the first step, the cybersecurity requirements for the pipeline system are decomposed into their constituent NFR softgoals. In the second step the architecture of the specific pipeline system is decomposed into its components and connections, and this creates a hierarchy of operationalizing softgoals that represent these architectural constituents. In the third step, a determination of the contributions made by the operationalizing softgoals to the NFR softgoals is made. In the fourth step, an analysis of the cybersecurity incident report is performed and cause softgoals are identified; the contributions made by these cause softgoals to operationalizing softgoals are also captured. In the fifth step, based on the cybersecurity body of knowledge, causes for these cause softgoals are identified, thereby creating a hierarchy of cause softgoals. In the sixth step, an association is made between the leaf cause softgoals and their corresponding operationalizing softgoals by means of tracebacks. In the seventh step, the propagation rules of

the NFR Approach are applied to determine those tracebacks that are satisficed and from which conclusions about the root causes for the cybersecurity incident can be made. In the last step, the eighth step, justifications are captured for all elements in the SIG by means of claim softgoals.

In a SIG represented graphically, the NFR softgoals and their decompositions are shown at the top of the figure (the requirements part), the operationalizing softgoals and their decompositions are shown in the middle of the figure (the design part), while the cause softgoals, their decompositions, and tracebacks are shown in the bottom of the figure (the cause analysis part).

4. Case Study of Florida Water Supply System Incident

In this section, the eight steps of the NFR Approach will be applied to analyze the root causes for a real incident - the Florida water supply system cyberattack [6, 16] that occurred in February, 2021.

4.1. Decomposition of Cybersecurity Requirements

Cybersecurity of the pipeline is the goal to be achieved by the design of the pipeline system. This goal is represented as an NFR softgoal in the Softgoal Interdependency Graph (SIG) of Figure 3 at the top by the light-bordered rectangle named *Cybersecurity of Pipeline*. This NFR softgoal is AND-decomposed into two child NFR softgoals: *Cybersecurity of Pumping Stations* and *Cybersecurity of Monitoring Stations*; the AND-decomposition is indicated by the single arc joining the two child NFR softgoals to the parent NFR softgoal, namely, *Cybersecurity of Pipeline*. The reason for the AND-decomposition is the fact that both child softgoals need to be satisficed for the parent softgoal to be satisficed; that is, the pumping stations and monitoring stations both need to be secure. The NFR softgoal *Cybersecurity of Pumping Stations* is further AND-decomposed into three NFR child softgoals: *Control System Security, Network Security*, and *Enterprise IT Security*; the reason for the AND-decomposition is that all three child softgoals need to be satisficed for the parent to be satisficed (that is, it is not sufficient for a subset of them to be satisficed for the parent to be satisficed). This is because, for the pumping stations to be secure their control system, the network connecting the control system to the pumping stations, and the enterprise IT that allows remote control should all be secure. Likewise the NFR softgoal *Cybersecurity of Monitoring Stations* is AND-decomposed into the three child softgoals *Network Security, Enterprise IT Security*, and *Acquisition System Security*; this means that monitoring stations' security can only be achieved if the network connecting them to the control system is secure, the enterprise IT that allows remote monitoring is secure, and the data acquisition system that collects the monitored information is secure. The top part of the SIG of Figure 3 has captured the high-level requirements for a secure pipeline and this completes the first step of the NFR Approach.

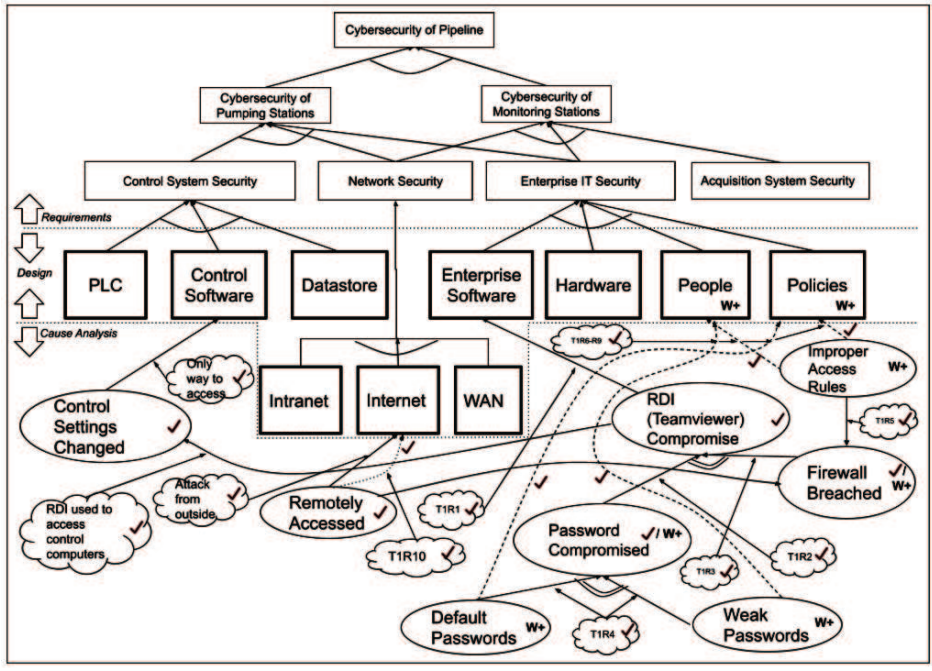

Figure 3. Softgoal interdependency graph for root cause analysis

4.2. Decomposition of Pipeline Architecture

In the middle part of the SIG of Figure 3, corresponding to the second step of the NFR Approach, system's design is captured by means of the operationalizing softgoals. The Programmable Logic Control (PLC) that actually controls the pumps at the pumping stations, the software that helps control these PLC's, and the datastore associated with logs for this control software, each impact the cybersecurity of the control system; each of these components are represented by operationalizing softgoals (rectangles with darker borders) in the SIG of Figure 3 with corresponding names of *PLC, Control Software*, and *Datastore*, respectively; the fact that they all are required components is captured by the AND-decomposition symbol (the single arc) near the NFR softgoal *(Control System Security)* they are related to. In a similar manner, the design components that impact network security are the intranet, the Internet, and the Wide Area Networks (WAN's) used to connect the distributed system components to the control system; and these are captured by their corresponding AND-decomposed (since all the three types of network connections are used in a modern pipeline control system) operationalizing softgoals with respective names *Intranet, Internet*, and *WAN*. The enterprise IT security is impacted by several design components including software, hardware, people that use the system, and policies for the system operation and use; each of these components are captured by their corresponding AND-decomposed operationalizing softgoals with names *Enterprise Software, Hardware, People*, and *Policies*, respectively.

4.3. Determine Contributions Between Softgoals

The contributions between the operationalizing softgoals and the NFR softgoals are indicated by the MAKE contributions that are involved in the AND decompositions discussed in Section 4.2; these contributions traverse the dotted line separating the requirements part of the SIG from the design part of the SIG. This completes the third step of the NFR Approach.

4.4. Identification of Cause Softgoals

Given the decomposition of both security requirements and the design components affecting security, in the fourth step of the NFR Approach process, cause analysis is performed using the SIG developed so far. From the incident reports [6, 16] it can be observed that the control software settings were changed and that this software was accessed using a Remote Desktop Interface (RDI) software, which in this case appears to be TeamViewer software [17]. This is captured in the SIG with the cause softgoal *Control Settings Changed* that points to *Control Software* operationalizing softgoal indicating that this cause affected the control software component; or, in the NFR Approach terminology, this cause strongly satisficed the operationalizing softgoal *Control Software*. *RDI (Teamviewer) Compromise* is the cause softgoal that affected the *Enterprise Software* softgoal allowing the compromise to occur; the arrow between the *RDI (Teamviewer) Compromise* cause softgoal and *Control Settings Changed* cause softgoal captures the fact that the latter caused the former (or that the latter strongly satisfices the former). The fact that this incident occurred from outside the organization using the Internet is captured by the cause softgoal *Remotely Accessed* that satisfices the operationalizing softgoal *Internet*. Cause softgoals determined so far are based on incident reports. These cause softgoals are marked with the satisficed symbol (the check marks) that indicates their validity since they are based on published reports.

4.5. Identification of Cause Softgoal Chains

In step five of the NFR Approach, in order to identify the causes for the known cause softgoals, the body of knowledge on cybersecurity is used to refine cause softgoals into child cause softgoals. It is known that RDI compromise can occur due to either the login credentials of an existing user being compromised or a successful connection being made to a valid user from the outside. Therefore, two possible causes can be identified for the RDI compromise: passwords being compromised and the firewall being breached; these two causes are captured by the cause softgoals *Password Compromised* and *Firewall Breached*. Since these two softgoals are two possibilities in an either-or situation (that is, either one or both could have caused the RDI compromise), this is indicated on the SIG using the double-arc that represents the OR-decomposition in the NFR Approach. Further, two causes can be identified for passwords being compromised: default passwords were used or weak passwords were used; they are captured by cause softgoals with names *Default Passwords* and *Weak Passwords*. Moreover, these cause softgoals

are in an OR-decomposition (indicated by the double arc) since either one or both may be the cause for the parent cause softgoal. Likewise the cause for firewall breach is the possible existence of improper access rules and this is captured by the cause softgoal *Improper Access Rules*.

4.6. Identification of Tracebacks

In step six of the NFR Approach, tracebacks are determined: traceback is the relationship between cause softgoals and design components and these are captured by the dashed arrows from cause softgoals to the design components. Attempt is made to traceback the bottom-most cause softgoals in the chain of causes to the design components so that the root causes reflect reality better. In this case study, the use of default passwords can be traced back to the fact that people did not change them and the use of weak passwords can be traced back to the password policies; these are captured by the traceback contributions from *Default Passwords* and *Weak Passwords* cause softgoals to the design components represented by the operationalizing softgoals *People* and *Policies*, respectively. Likewise, the cause softgoal *Improper Access Rules* can be traced back to the people who set them as well as to the policies that allowed them to set these rules and these are captured by the two tracebacks from this cause softgoal. Also *Remotely Accessed* cause softgoal is directly traced back to its parent operationalizing softgoal, namely, *Internet*.

4.7. Application of Propagation Rules

In the seventh step of the NFR Approach, the propagation rules are applied and root causes are identified. First an assumption can be made that all claim softgoals, that is, the rationale for elements in the SIG are correct - the justification for this is that the claims were derived from available information and they are assumed to be correct (this point is discussed later). Therefore, by rule R1 all claim softgoals are satisficed. Because of this, by propagation rule R2, all contributions in the SIG are also satisficed since all contributions in the requirements and design part are all MAKE-contributions (this point is discussed later). In the cause analysis part (bottom part) of the SIG in Figure 3, all claim softgoals that are satisficed are indicated by check marks. Also the cause softgoals *Control Settings Changed, Remotely Accessed*, and *RDI (Teamviewer) Compromise* are all satisficed since these were obtained from incident reports - the check marks inside these softgoals indicate this satisficing. At this point the labels for the cause softgoals *Password Compromised* and *Firewall Breached* are not known; however, by propagation rule R10, since the parent cause softgoal *RDI (Teamviewer) Compromise* is satisficed, these two child softgoals involved in OR-decomposition must be have one of the two possibilities: both are satisficed (check marks apply for both) or either one is satisficed (check mark for one and W+ for the other). In the second case, the non-satisficed softgoal gets the weakly satisficed symbols (W+) since the information is based on the known body of knowledge but not yet known to be a fact. In a similar manner by applying propagation rule R10 the

child cause softgoals *Default Passwords, Weak Passwords*, and *Improper Access Rules* get the W+ label - the label will change to satisficed the moment the cause is verified. Moreover, by propagation rule R2, all tracebacks are satisficed (since the claims justifying them are satisficed by rule R1) and these are indicated by check marks in the SIG; therefore, by rule R6, all tracebacks propagate the labels of children to their parent. That is, the operationalizing softgoals *People* and *Policies* both get weakly satisficed labels (W+) while the operationalizing softgoal *Internet* gets the satisficed label (check mark).

4.7.1. Identification of RootCauses for the Incident

There are five tracebacks in the SIG of Figure 3 and they point to the operationalizing softgoals of *People, Policies*, and *Internet*, which correspond, respectively, to the three components that affect the system's cybersecurity: people in the organization, policies followed by the organization, and the Internet connection. These tracebacks inform us that these system components are the root causes for this cybersecurity incident as well. The *People* component of the system indicates that better cybersecurity training and system use training are required for people using the system, the *Policies* component indicates that better policies are needed for setting passwords and firewall rules, and the *Internet* component tells us that this network requires better protection at the Internet side for preventing such incidents from occurring again. These conclusions are captured in Table 2.

4.8. Capturing Rationale Using Claim Softgoals

In the last step, the eighth step of the NFR Approach, the rationale for each element in the SIG is captured by means of claim softgoals indicated by cloud-shaped figures in the SIG. While claim softgoals may be attached to any softgoal, contribution, traceback, and decomposition in the SIG, for simplicity sake, the claims have been shown only for the contributions and tracebacks in Figure 3. The cause softgoal *Control Settings Changed* satisfices the operationalizing softgoal *Control Software* due to the reason "Only way to access" which is captured by the claim softgoal with the name of the reason. The cause softgoal *RDI (Teamviewer) Compromise* satisfices the cause softgoal *Control Settings Changed* due to the claim softgoal "RDI used to access control computers". The cause softgoal *Remotely Accessed* satisfices the operationalizing softgoal *Internet* due to "Attack from outside" justification. Other claim softgoals are indicated in Table 1 and their corresponding row numbers are shown in the claim softgoals in the SIG of Figure 3; for example, "T1R1" means Table 1 Row 1, and so on. Five rows, 6 through 10, of Table 1 capture the traceback justifications in the SIG and this is shown in Figure 3 by "T1R6-R9" claim softgoal and by "T1R10" claim softgoal. All claim softgoals are marked as satisficed (check symbols) since, based on current knowledge, these claims are known to be true.
From Weak Passwords cause People Insufficient cybersecurity training and equipment softgoal to People use training for users. operationalizing softgoal From Improper Access Rules People Insufficient technology training provided to IT cause

Table 1. Rationale for Contributions and Tracebacks in the SIG of Figure 3

Row No.	SIG Element	Source Softgoal	Destination Softgoal	Claim
1	Contribution	RDI (Teamviewer) Compromise	Enterprise Software	"RDI (Teamviewer) is part of the enterprise software used for remote access"
2	Contribution	Password Compromised	RDI (Teamviewer) Compromise	"If a user's password is compromised then the RDI is compromised"
3	Contribution	Firewall Breached	RDI (Teamviewer) Compromise	"If the firewall is breached then any actor from outside can compromise RDI"
4	Contribution	Default Passwords and Weak Passwords	Password Compromised	"Default passwords or weak passwords enable password compromise"
5	Contribution	Firewall Breached	Improper Access Rules	"Improper access rules permit firewall to be breached"
6	Traceback	Weak Passwords	People	"People (employees) set weakpasswords"
7	Traceback	Improper Access Rules	People	"IT technicians set the improper firewall rules"
8	Traceback	Default Passwords	Policies	"IPolicies allow employees to continue using default passwords without changing them"
9	Traceback	Improper Access Rules	Policies	"Policies allow improper access rules to be set in the firewall"
10	Traceback	Remotely Accessed	Internet	"Internet access allowed this incident to occur"

softgoal to People personnel who set firewall access rules. operationalizing softgoal From Default Passwords cause Policies Policies on changing default passwords by softgoal to Policies users not clear or not enforced. operationalizing softgoal From Improper Access Rules Policies Policies for setting or updating firewall access by cause softgoal to Policies IT personnel not clear or not enforced. operationalizing softgoal From Remotely Accessed cause Internet Internet accesses need better protection. softgoal to Internet operationalizing softgoal

5. Discussion of Results

The NFR Approach provides a systematic approach for detecting root causes for cybersecurity incidents involving pipelines used in the critical infrastructure. By dividing the process of detecting root causes into steps, the security analyst gets sufficient opportunity to ensure that all pertinent facts are captured

Table 2. Root Causes Identified Using the SIG of Figure 3

Traceback	Design Component	Root Cause
From *Weak Passwords* cause softgoal to *People* operationalizing softgoal	People	Insufficient cybersecurity training and equipment use training for users
From *Improper Access Rules* cause softgoal to *People* operationalizing softgoal	People	Insufficient technology training provided to IT personnel who set firewall access rules
From *Default Passwords* cause softgoal to *Policies* operationalizing softgoal	Policies	Policies on changing default passwords by users not clear or not enforced
From *Improper Access Rules* cause softgoal to *Policies* operationalizing softgoal	Policies	Policies for setting or updating firewall access by IT personnel not clear or not enforced
From *Remotely Accessed* cause softgoal to *Internet* operationalizing softgoal	Internet	Internet accesses need better protection

in the SIG and that the subsequent cause analysis is based on the known facts. As discussed above, in the case study involving the Florida water system attack that occurred in February of 2021, the NFR Approach helps determine that the root causes can be traced to system issues that are usually under control of entities affected by the attack: user training, IT personnel training, and secure Internet access.

The three parts of the SIG make it easier for the analyst to recognize those system issues that primarily affect the cybersecurity requirements; for example, having high performance machines will definitely help user experience but are not directly related to cybersecurity requirements. The SIG allows the cybersecurity requirements to be the focus of analysis and consider only those aspects of the system design that directly affect cybersecurity requirements. Therefore, the subsequent cause analysis phase allows determination of root causes that are tied to those system components specifically related to cybersecurity requirements; in this way, it can be ensured that causes identified are traceable to cybersecurity requirements for the pipeline system.

In a typical application of the NFR Approach, cause softgoals are usually not required. Here, the concept of cause softgoals was introduced to systematically identify root causes for a cybersecurity incident. Moreover, in the NFR Approach, the propagation rules are applied to identify the extent of satisficing of the root NFR softgoals - the NFR softgoals at the top of the SIG; here, the propagation rules are applied to identify the extent of satisficing of the leaf operationalizing softgoals since the goal is to identify the root causes originating in the design of the system. Another point to note is that all softgoals have been named without their *Topic* which was discussed earlier - adding *Topic's* to softgoal names will

make them more appropriate to the specific system: for example, in the SIG of Figure 3, the PLC can be explicitly referred to as the one at the pumping station in Dallas, for example, as *PLC [Dallas]*, to differentiate that specific PLC from others along the pipeline. Also, criticality symbols for softgoals have not been used; these symbols can significantly make the analysis more specific and accurate by allowing the use of priorities for softgoals for the system under investigation.

In the SIG of Figure 3, the *Acquisition System Security* NFR softgoal was not decomposed further since this acquisition system was considered not relevant to this incident; however, if that is not the case, this NFR softgoal can be further decomposed as needed. Moreover, in the SIG of Figure 3, it was assumed that the design components consisted of one layer only; that need not be the case and, in fact, more refinements of the design will help identify the root causes better. For example, the design decomposition (or operationalizing softgoal decomposition hierarchy) of Figure 4 can help identify root causes more clearly by assigning tracebacks to their appropriate and more specific design component. The decomposition in Figure 4 is just the design part of the SIG of Figure 3 and shows how the *People* softgoal of Figure 3 may be decomposed further to get more accurate information during cause analysis; here the tracebacks from the two cause softgoals shown in Figure 3 (*Improper Access Rules* and *Default Passwords*) go to design components *Firewall Training* for IT personnel and *Cybersecurity Training* for operations personnel.

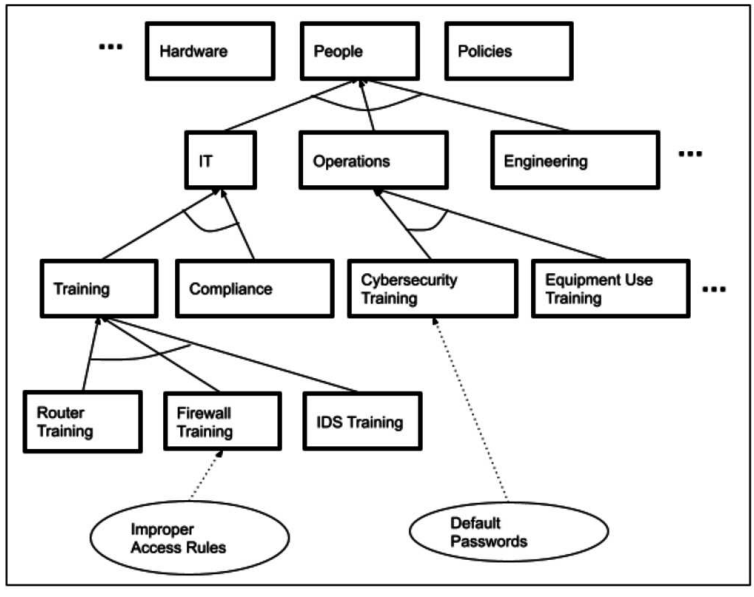

Figure 4. Detailed decomposition of operationalizing softgoals

To decompose cause softgoals obtained from incident reports the body of available cybersecurity knowledge was used; in addition, other sources such as

organizational experience, consultants, and industry best practices may also be used to identify the chain of causes. As long as the reasons for these decompositions are captured by claim softgoals, there will always be a historical record of decision-making. In the future, if new evidence comes to light then the SIG can be updated accordingly and more appropriate root causes may be identified. An important point to note is that the eight steps in the application of the NFR Approach are not just sequential but also iterative; it is possible to go back to any prior step when more knowledge becomes available and restart the analysis from there. This also applies to the assumption made earlier in Section 4.7 when it was assumed that all claim softgoals were satisficed - if during subsequent investigation it turns out that some or all of the current claims are unfounded then those claims will become denied and a restart of SIG analysis will be required with new claims added to deny earlier ones and support newer ones.

In the SIG of Figure 3, the requirements of the system were decomposed in a specific manner; however, this decomposition is not necessarily unique and can be done in a manner most appropriate to the system under analysis. For example, it is possible to decompose each pumping station and each monitoring station separately; however, here simpler version has been used for illustrating the NFR Approach to analyze causes of cybersecurity incidents on pipelines. Moreover, the requirements may interact with each other as well either synergistically or in a conflicting manner: for example, control from anywhere and anytime requires that the control system be exposed to the Internet but this impacts security negatively since potential attackers may see a vulnerability. As another example, measures to improve reliability such as using wired connections also improves security since wired connections are more difficult to eavesdrop. But the SIG can capture these synergistic and conflicting interactions by the use of contributions, decompositions, and propagation rules.

In the application of the NFR Approach to the Florida water system cyberattack it was stated that since the traceback points to the People operationalizing softgoal it most likely means that employees need better cybersecurity and equipment-use training; however, this is not necessarily the only possibility - employees may be well-trained but the process of using the system may be the reason for the apparent error or there may have been an social engineering vulnerability. In these cases it is possible to refine the SIG to pinpoint the exact cause for the failure in the People component of the system, for example, as shown in Figure 4.

Finally, as can be seen, the SIG also serves as the historical record of all system related information as far as cybersecurity incident analyses are concerned. As more information becomes available the SIG becomes more accurately populated and the resulting analyses become more useful. Every version of the SIG can be saved in an information system to save the progress of the analysis as well as for any future audits. For example, an assumption was made that there exist MAKE contributions between operationalizing softgoals and NFR softgoals and between NFR softgoals themselves - if this assumption turns out to be

false later, then those contributions will need to be changed to HELP, HURT, or BREAK, but these changes will be made with accompanying claims so that historical record during the investigative analysis is preserved.

6. Conclusion

Pipelines carrying critical fluids including water, oil, gas, and other chemicals are part of the critical infrastructure of the nation. Recent cybersecurity incidents on pipeline systems indicate the extent to which economic damage can be done by malware and malicious actors. While cybersecurity protection is important, what may be even more important, especially after a cybersecurity incident, is to identify root causes so that a repeat of the incident is prevented. Very often cybersecurity incidents are the result of human error and so pure technical defenses for cybersecurity may not be sufficient; however, unless the root causes are identified to be related to human errors wrong measures may be taken to avoid repetition of the incident. While several techniques exist to identify root causes including the five whys, fish-bone diagrams, and causal factors analysis, in this paper, the NFR Approach is presented as an alternative technique for identifying root causes after a cybersecurity incident on pipelines.

The NFR Approach is a goal-oriented approach that considers the primary causes as the goal to be achieved by further examination of evidence and body of knowledge. By refining the primary causes, the NFR Approach identifies the root causes in a systematic manner. To this end the NFR Approach uses a softgoal hierarchy that consists of requirements softgoals, design softgoals, cause softgoals, their decompositions, contributions, tracebacks, and rationale for all elements in the hierarchy captured in the form of claim softgoals. The resulting graphical hierarchy, called the Softgoal Interdependency Graph (SIG), captures all decisions taken during cause analysis for historical record keeping. There are eight steps in the NFR Approach which when applied sequentially and iteratively will help identify root causes for cybersecurity incidents on pipelines.

The NFR Approach was applied to a real cybersecurity incident that took place in Florida in February, 2021, on a water supply system and identified the potential root causes for the incident. During the application of the NFR Approach, the SIG was first created based on available information of the system and the incident, then contributions and tracebacks were identified, subsequently propagation rules of the NFR Approach were applied, and finally the root causes for the incident were determined from the SIG: in this case, the root causes included insufficient cybersecurity training for the control system operators and poor enforcement of password policies for users.

In the future, the NFR Approach can be applied to analyze root causes of cybersecurity incidents on other pipeline systems. Also, the NFR Approach can be used to identify root causes for cybersecurity incidents on other systems besides just pipelines such as root cause identification for cybersecurity incidents on cyber physical systems in general. Another line of future research is to develop

a spreadsheet template that can be used by practitioners to quickly apply the different steps of the NFR Approach for their systems of interest and identify the root causes of any cybersecurity incidents on those systems.

References

[1] M Yoon, C B Warren and S Adam, 2007, *Pipeline System Automation and Control*, ASME Press

[2] *Framework for Improving Critical Infrastructure Cybersecurity, Version 1.1*, National Institute of Standards and Technology, April 2018 https://nvlpubs.nist.gov/nistpubs/CSWP/NIST.CSWP.04162018.pdf

[3] Infrastructure Security, Cybersecurity and Infrastructure Security Agency, https://www.cisa.gov/infrastructure-security

[4] N Subramanian, 2008, *Improving Security of Oil Pipeline SCADA Systems Using Service-Oriented Architectures*, Lecture Notes in Computer Science, November, **5333**, 344 – 353

[5] Solutions for Oil & Gas Pipelines, Brochure, Honeywell, May 2016 https://www.honeywellprocess.com/library/marketing/brochures/Solutions-OilandGas-pipelines.pdf

[6] M S Warren, Florida Water System Hack Offers Lessons for Other States, April 19 https://www.govtech.com/security/florida-water-system-hack-offers-lessons-for-other-states.html

[7] K Lyon, June 5, 2021, Hackers reportedly used a compromised password in Colonial Pipeline cyberattack https://www.theverge.com/2021/6/5/22520297/compromised-password-reportedly-allowed-hackers-colonial-pipeline-cyberattack

[8] B Miller and D C Rowe, October 2012, *A Survey of SCADA and Critical Infrastructure Incidents*, Proceedings of the ACM Special Interest Group on Information Technology Education, Calgary, Canada

[9] *What is Cybersecurity?*, Cybersecurity and Infrastructure Security Agency, November 14, 2019, https://us-cert.cisa.gov/ncas/tips/ST04-001

[10] M Ahola, *The Role of Human Error in Successful Cyber Security Breaches*, https://blog.usecure.io/the-role-of-human-error-in-successful-cyber-security-breaches

[11] Root Cause Analysis https://des.wa.gov/services/risk-management/about-risk-management/enterprise-risk-management/root-cause-analysis

[12] N Subramanian, 2021, *Root Cause Analysis*, Encyclopedia of Cryptography and Security (3rd Ed.), Editors: S. Jajodia and H. C. A. van Tilborg, Springer Publication https://link.springer.com/referenceworkentry/10.1007/978-3-642-27739-9_1498-1

[13] V C Moreno et. al., May 2018, *Analysis of Physical and Cyber Security-Related Events in the Chemical and Process Industry*, Journal of Process Safety and Environmental Protection, Elsevier, **116** 621 – 631

[14] M Panini et. al., May 2021, *Analysis of Cybersecurity-related Incidents in the Process Industry*, Journal of Reliability Engineering System Safety, Elsevier, **209**

[15] N Subramanian and J Zalewski, June 2016, *Quantitative Assessment of Safety and Security of System Architectures for Cyberphysical Systems Using the NFR Approach*, IEEE Systems Journal, **10** (2) 397 – 409

[16] B Krebs, February 10, 2021, What's most interesting about the Florida water system hack? That we heard about it at all. https://krebsonsecurity.com/2021/02/whats-most-interesting-about-the-florida-water-system-hack-that-we-heard-about-it-at-all/

[17] TeamViewer company website https://www.teamviewer.com/en-us/

Nary Subramanian is an Associate Professor of Computer Science at the University of Texas at Tyler. He is also the Co-Founder and the Chief Technology Officer of dMACQ Software Pvt. Ltd., Mumbai, India. He obtained his Ph.D. in Computer Science from the University of Texas at Dallas. He served in the industry for fifteen years in engineering, sales, and management before joining academia. He co-founded and successfully organized ten editions of the International Workshop on System/Software Architectures, established and served as the Director of the Center for Petroleum Security Research at UT Tyler, and received a fellowship from Air Force Research Lab, Rome, NY. His research interests include non-functional requirements, cyber-physical systems, and cybersecurity.